

A DICTIONARY-BASED COMPARISON OF RUSSIAN AND
UZBEK SET EXPRESSIONS

Tagaeva Feruza Erkinovna

Lecturer

Department of Russian Language and Literature

Bukhara State University

Uzbekistan, Bukhara

Abstract. The article compares Russian and Uzbek set expressions on the basis of three phraseological dictionaries, the Uzbek dictionary by Sh. Rahmatullayev and co-authors (3,938 units) and two Russian dictionaries of different types, the academic dictionary edited by A.I. Fedorov (6,884 units) and the general dictionary by Yu.A. Larionova (2,855 units). The units were lemmatised and classified by their thematic components, and the differences were tested with the χ^2 test. Somatic components occur in 24.3% of Uzbek units against 17.3% and 20.9% in the Russian dictionaries, and components denoting the inner world of a person occur in 5.5% against 2.4% and 2.5%. Animal and colour components show no significant differences. The somatic profiles of the two languages are similar ($p = 0.70$ and 0.71), but Uzbek prefers the eye, the mouth, the face and the liver, while Russian more often uses the shoulder and the throat. Uzbek locates feelings in *ko'ngil*, *jon* and *bag'ir*, where Russian uses *dyua* and *cepdye*. Twelve verified pairs illustrate four types of correspondence, from full equivalents such as *совать нос* and *burnini suqmoq* to units without a somatic counterpart such as *ko'zi to'q*.

Keywords. set expression, phraseological unit, Russian, Uzbek, phraseological dictionary, somatism, comparative phraseology, equivalence.

INTRODUCTION

Set expressions are the part of vocabulary in which the experience and the values of a people are fixed most directly. Russian phraseology developed a detailed theory of such units, from the classification of V.V. Vinogradov [1] and the structural description of N.M. Shansky [7] to the cultural approach of V.N. Telia, who treated idioms as signs of a national mentality [4]. Comparative phraseology, represented by V.M. Mokienko for the Slavic languages [3] and by D. Dobrovol'skij and E. Piirainen for a wide range of European and Asian languages [8], showed that idioms built on the names of body parts, animals, numbers and colours are the most suitable material for comparison, because their components are found in every language.

Uzbek phraseology was established as a field by Sh. Rahmatullayev, whose explanatory phraseological dictionary [10] became the basis of later lexicographic work, and its revised version of 2022 is now the most complete description of Uzbek idioms [11]. Russian phraseography offers several dictionaries of different types, from the classical dictionary edited by A.I. Molotkov [6] to the large academic collection edited by A.I. Fedorov [5] and modern general dictionaries [2]. Russian and Uzbek idioms are usually compared on selected examples, and quantitative comparisons of whole dictionaries remain rare.

The aim of the study is to compare the component composition and some structural features of Russian and Uzbek set expressions on the basis of three phraseological dictionaries, to identify the differences that remain stable across dictionaries of different types and to illustrate the main types of interlingual correspondence.

MATERIALS AND METHODS

The Uzbek material is the phraseological dictionary by Sh. Rahmatullayev, N. Mahmudov, Z. Xolmanova, I. O‘razova and K. Rixsiyeva [11], used in an open machine-readable version that contains 4,408 meanings of 3,938 units. 363 of them are historical units marked as such in the dictionary. The Russian material comes from two dictionaries of different types. The dictionary edited by A.I. Fedorov [5] contains about 7,000 units of the literary language of the late 18th to 20th centuries, many of which are absent from earlier dictionaries. The dictionary by Yu.A. Larionova [2] is a general dictionary of common modern idioms. Headwords were extracted automatically from the recognised text of scanned editions, variants and cross-references were removed and duplicates were merged, which gave 6,884 and 2,855 units.

Russian units were lemmatised with pymorphy [9]. Uzbek units were analysed with stem patterns that allow possessive and case endings and the alternations of the type *yurak* and *yuragi* or *og‘iz* and *og‘zi*, and homonymous forms were checked manually, for example *bosh* «head» and *boshqa* «other», *ot* «horse» and *oti* «his name», *yuz* «face» and «hundred», *uchi* «its tip». Six thematic classes were distinguished, namely somatic components (more than 40 body parts), components of the inner world (*dyua*, *dyx*, *ko‘ngil*, *dil*, *jon*, *ruh*), numerals, religious and mythological components, animal names and colour terms. A unit was counted as verbal if it contains an infinitive in Russian or a form in *-moq* or *-maslik* in Uzbek. Differences between dictionaries were tested with the χ^2 test for 2×2 tables, and the somatic profiles were compared with Spearman’s rank correlation for body parts found in at least 20 units. Pairs of equivalents were selected manually and verified by the definitions in both dictionaries.

RESULTS

3.1 Structure of the units

Table 1 gives the main parameters of the three dictionaries. Uzbek units are shorter. Two-word units make up 63.2% of the Uzbek inventory against 53.2% in the dictionary edited by A.I. Fedorov, while the mean length is 2.56 and 2.65 words. The citation forms in the dictionary by Yu.A. Larionova often include typical partners of the idiom, so its length was not compared. Verbal units predominate in Uzbek (58.2%) and make up less than half of the Russian inventory in both dictionaries ($p < 0.001$).

Table 1. Parameters of the dictionaries

Parameter	Russian, Fedorov	Russian, Larionova	Uzbek
Units	6,884	2,855	3,938
Two-word units, %	53.2	n/a	63.2
Three-word units, %	33.5	n/a	23.0
Four or more words, %	13.3	n/a	13.1
Mean length, words	2.65	n/a	2.56
Verbal units, %	40.7	48.3	58.2
Somatic units, %	17.3	20.9	24.3

3.2 Thematic components

Figure 1 compares the six thematic classes. Somatic components are the largest class in all dictionaries, but they are more frequent in Uzbek (24.3%) than in both Russian dictionaries (17.3% and 20.9%, $p < 0.001$ and $p = 0.001$).

Components of the inner world are twice as frequent in Uzbek (5.5% against 2.4% and 2.5%, $p < 0.001$ in both comparisons). Animal names and colour terms occur with almost the same frequency in the three dictionaries, and the differences are not significant. Numerals and religious components differ only from the dictionary edited by A.I. Fedorov, which contains many obsolete units with *боғ*, *чëpm* and *бес*, so these differences depend on the type of dictionary.

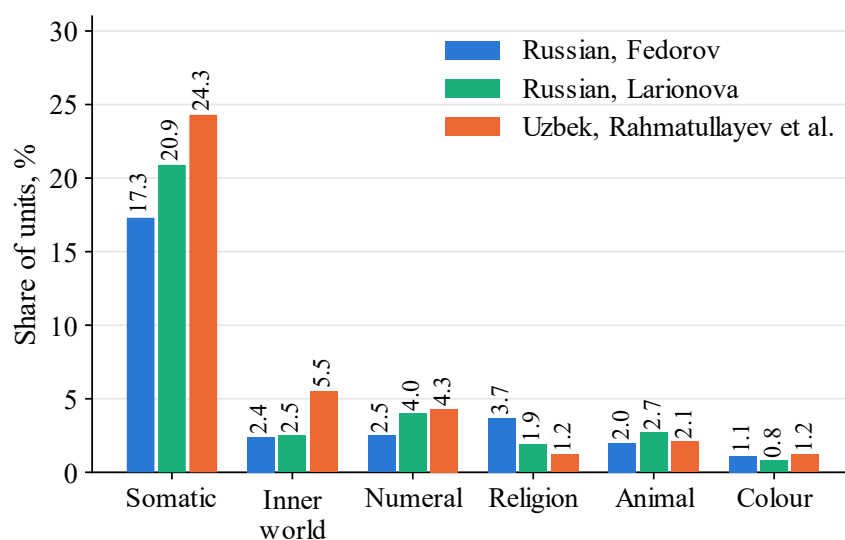


Figure 1. Thematic components in Russian and Uzbek set expressions.

3.3 Somatic profile

Figure 2 shows how often the fifteen most productive body parts occur in each dictionary. The profiles of the two languages are similar, and the rank correlation is high ($\rho = 0.70$ with the first Russian dictionary and $\rho = 0.71$ with the second, $p < 0.01$). The eye, the hand and the head lead in all dictionaries. Against this common background Uzbek shows clear preferences. The eye (*ko‘z*) occurs in 42.9 units per thousand against 23.4 and 28.4 in Russian, the mouth (*og‘iz*) in 20.8 against 5.1 and 6.3, and the face (*yuz*, *bet*, *chehra*) in 19.3 against 2.6 and 6.7. The liver, which is almost absent from Russian

idioms, is found in 27 Uzbek units. Russian more often uses the shoulder and the throat ($p < 0.05$ in both dictionaries).

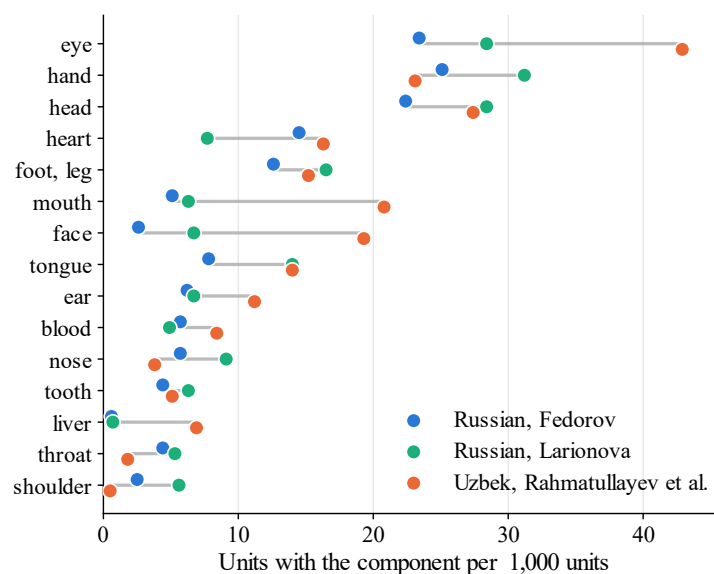


Figure 2. Body parts in Russian and Uzbek set expressions.

3.4 Inner world and the seats of feelings

Figure 3 compares the words that locate feelings and character. In Russian this role belongs to *душа* and *сердце*, while *дух* and *печень* are rare. In Uzbek the leading component is *ko'ngil* (with *dil* 32.8 units per thousand), followed by *jon* and *ruh* (21.8) and *yurak* (16.3). The liver in the form of *bag'ir* and *jigar* is used in 6.9 units per thousand, mostly for kindness, grief and cruelty, as in *bag'ri keng* «generous», *bag'ri qon* «grieving» and *bag'ri tosh* «heartless». Thus the Uzbek inventory distributes feelings among more organs than the Russian one.

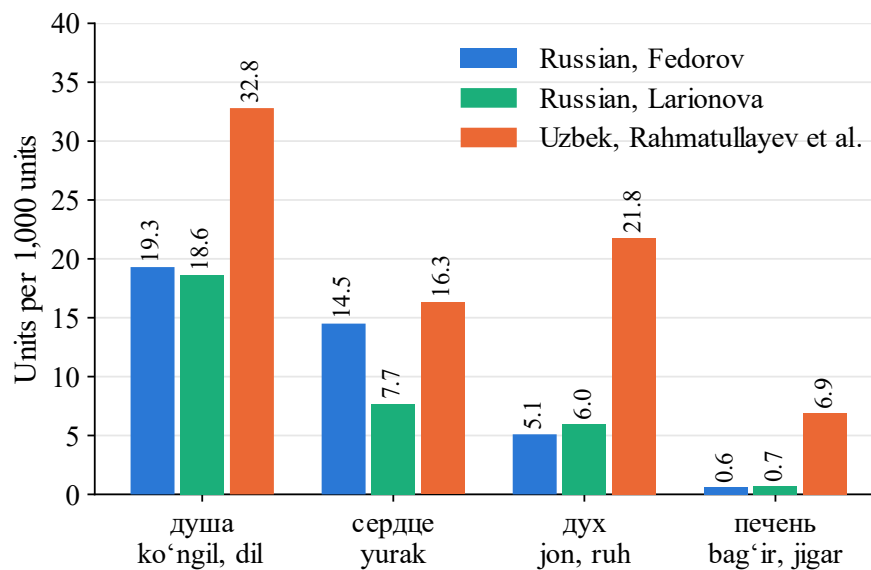


Figure 3. Words for the seats of feelings in Russian and Uzbek set expressions.

3.5 Types of correspondence

Table 2 gives twelve pairs verified in the dictionaries. Full equivalents share the component, the image and the meaning, as *совать нос* and *burnini suqmoq*. Partial equivalents have the same meaning but differ in the image, as *золотые руки* «gold hands» and *qo'li gul* «hands are a flower», or in the organ, as *каменное сердце* and *bag'ri tosh* «the liver is a stone». In *водить за нос* and *burnidan ip o'tkazib olmoq* the image of leading by the nose is common, but the meanings differ, deception in Russian and submission in Uzbek. Finally, *ko'zi to'q* «his eye is full», that is, not greedy, has no somatic counterpart in the Russian dictionaries.

Table 2. Types of Russian and Uzbek correspondences

Type	Russian	Uzbek	Meaning
Full	<i>совать нос</i>	<i>burnini suqmoq</i>	interfere
Full	<i>задирать нос</i>	<i>burni ko 'tarilmoq</i>	become arrogant
Full	<i>бросаться в глаза</i>	<i>ko 'zga tashlanmoq</i>	be conspicuous
Full	<i>закрывать глаза</i>	<i>ko 'z yummoq</i>	ignore deliberately
Full	<i>вертится на кончике языка</i>	<i>tilining uchida turmoq</i>	cannot recall a known word
Full	<i>рот до ушей</i>	<i>og 'zi qulog 'ida</i>	very happy
Partial	<i>ломать голову</i>	<i>bosh qotirmoq</i>	think hard
Partial	<i>золотые руки</i>	<i>qo 'li gul</i>	skilful
Partial	<i>каменное сердце</i>	<i>bag 'ri tosh</i>	heartless
Partial	<i>падать лицом в грязь</i>	<i>yuzi qora</i>	lose face
Image only	<i>водить за нос</i>	<i>burnidan ip o 'tkazib olmoq</i>	deceive / make obedient
None		<i>ko 'zi to 'q</i>	not greedy

DISCUSSION

The comparison confirms that Russian and Uzbek phraseology rely on the same basic set of body parts, which agrees with the cross-linguistic data of D. Dobrovol'skij and E. Piirainen [8]. The eye, the hand and the head are

the most productive components in both languages, and the pairs in Table 2 show that common images often lead to full or partial equivalents. This common basis follows from the universal experience of the human body and makes somatic units the most convenient material for teaching Russian idioms to Uzbek students.

The differences concern the degree of somatisation and the organs chosen for particular meanings. Uzbek uses body parts and words of the inner world more often, and it expresses social evaluation through the face and the mouth, as in *yuzi qora* «disgraced» or *og‘zi qulog‘ida* «very happy». The role of *ko‘ngil*, *jon* and *bag‘ir* corresponds to the cultural models of internal organs described for several Asian languages, where organs such as the liver serve as seats of feelings [12]. In Russian the same function is concentrated in *dyua* and *cepdye*, which carry rich cultural connotations in Russian idioms [4].

The use of two Russian dictionaries shows which results are stable. The higher somatisation of Uzbek, the frequency of components of the inner world and the preferences for the eye, the mouth, the face and the liver are confirmed by both Russian sources, while the differences in numerals and religious components depend on the dictionary. The study has limitations. The Russian data were obtained from recognised text, and the Uzbek analysis relies on stem patterns. The components were counted formally, without regard to their degree of motivation.

CONCLUSION

The study leads to four conclusions. First, somatic components are the largest thematic class of set expressions in both languages, but they are more frequent in Uzbek (24.3%) than in Russian (17.3% and 20.9%). Second, the somatic profiles of the two languages are similar, while Uzbek prefers the eye, the mouth, the face and the liver, and Russian the shoulder and the throat.

Third, Uzbek distributes feelings among *ko'ngil*, *jon*, *yurak* and *bag'ir*, whereas Russian concentrates them in *duша* and *сердце*. Fourth, Russian and Uzbek idioms show full equivalents, partial equivalents, units with a common image but different meanings and units without equivalents. The results can be used in bilingual phraseography and in teaching Russian to Uzbek students.

REFERENCES

1. Виноградов В.В. Об основных типах фразеологических единиц в русском языке // Виноградов В.В. Избранные труды. Лексикология и лексикография. М.: Наука, 1977. С. 140-161.
2. Ларионова Ю.А. Фразеологический словарь современного русского языка. М.: Аделант, 2014. 512 с.
3. Мокиенко В.М. Славянская фразеология. 2-е изд., испр. и доп. М.: Высшая школа, 1989. 287 с.
4. Телия В.Н. Русская фразеология. Семантический, прагматический и лингвокультурологический аспекты. М.: Языки русской культуры, 1996. 288 с.
5. Фразеологический словарь русского литературного языка конца XVIII–XX в.: в 2 т. / под ред. А.И. Фёдорова. Новосибирск: Наука, 1991. Т. 1. 340 с.; Т. 2. 274 с.
6. Фразеологический словарь русского языка / под ред. А.И. Молоткова. М.: Советская энциклопедия, 1967. 543 с.
7. Шанский Н.М. Фразеология современного русского языка. М.: Высшая школа, 1985. 160 с.
8. Dobrovol'skij D., Piirainen E. Figurative Language. Cross-Cultural and Cross-Linguistic Perspectives. Oxford: Elsevier, 2005. 419 p.

9. Korobov M. Morphological analyzer and generator for Russian and Ukrainian languages // Analysis of Images, Social Networks and Texts. AIST 2015. Communications in Computer and Information Science. Vol. 542. Cham: Springer, 2015. P. 320-332.
10. Rahmatullayev Sh. O'zbek tilining izohli frazeologik lug'ati. Toshkent: O'qituvchi, 1978. 407 b.
11. Rahmatullayev Sh., Mahmudov N., Xolmanova Z., O'razova I., Rixsiyeva K. O'zbek tili frazeologik lug'ati. Toshkent: G'afur G'ulom nomidagi nashriyot-matbaa ijodiy uyi, 2022. 636 b.
12. Sharifian F., Dirven R., Yu N., Niemeier S. (eds.) Culture, Body, and Language. Conceptualizations of Internal Body Organs across Cultures and Languages. Berlin; New York: Mouton de Gruyter, 2008. 431 p.